Lecture Friday 21 August

Linear Regression

- $\boxed{\text{Data set}}$ (y, x)

output / Target: y; input: x

assumption

$$y(x) = f(x) + N(0, \sigma^2)$$

$f(x)$: unknown

$N(0, \sigma^2)$: PDF ~ Normal distribution

$\mu = 0$

variance σ^2

Assume we have a known PDF, $p(x)$ (continuous)

Mean value / Expected value

$$\mu = \int dx\, x\, p(x)$$

Our case is that $p(x)$ is often unknown. ML is a frequentist approach.

$$(y, x) \longrightarrow \begin{aligned} x &= \{x_0, x_1, \ldots, x_{n-1}\} \\ y &= \{y_0, y_1, \ldots, y_{n-1}\} \end{aligned}$$

The mean value we estimate

$$\mu_y = \frac{1}{n}\sum_{i=0}^{n-1} y_i$$

$\mu_y \neq \mu$ (Exact value if $p(y)$ is known)

<u>sample mean.</u>

$$\sigma^2 = \int (x-\mu)^2\, p(x)\, dx$$

sample variance

$$\sigma_x^2 = \frac{1}{n}\sum_{i=0}^{n-1} (x_i - \mu_x)^2$$

$$\sigma_x^2 \neq \sigma^2$$

$$\boxed{f(x) \simeq \tilde{y}(x)}$$

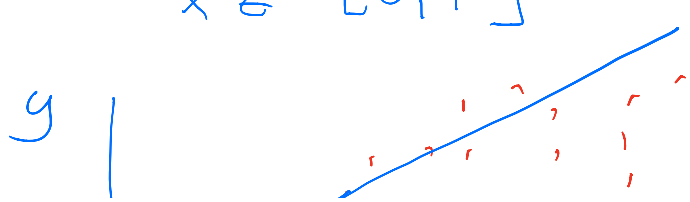
$$y(x) \rightarrow y(x_i) = y_i$$

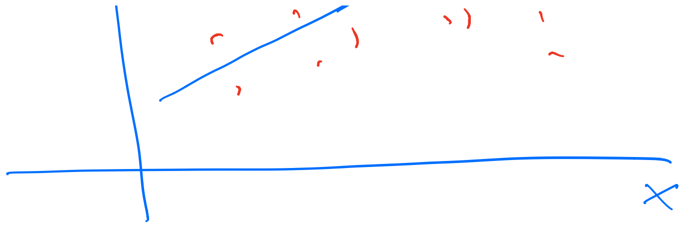
$$\boxed{y(x_i) \simeq \tilde{y}_i + \varepsilon_i}$$

$$\varepsilon_i \sim N(0, \sigma^2)$$

- $\boxed{\text{Model}}$

$$x \in [0,1]$$





$$\tilde{y}_i = \tilde{y}_i(x_i) = \beta_0 + \beta_1 x_i$$

$$\beta = \{\beta_0, \beta_1\}$$

How can we find β so that the difference between y (target) and $\tilde{y}$ is as small as possible?

$\boxed{\text{Cost/error/loss function}}$

Relative error $\left|\dfrac{y-\tilde{y}}{y}\right|$

$$y-\tilde{y} = \sum_{i=0}^{n-1}(y_i-\tilde{y}_i)$$

MSE = Mean squared error

$$= \frac{1}{n}\sum_{i=0}^{n-1}(y_i-\tilde{y}_i)^2$$

$$= C(\beta|x) \quad (= C(\beta))$$

$\hat{\beta} = \quad C(\beta|x)$

$$\beta = \underset{\beta}{\operatorname{argmin}} \; L(\beta|X)$$

optimal

Data set (y, x)

$$X \in \mathbb{R}^{n \times p}$$

features/predictor

data entries

Feature matrix.

$n = 100$, patient

$p = 10$, specific features of a given tumour

$$X = \begin{array}{c} \\ n=0 \\ n=1 \\ \\ \vdots \\ n=99 \end{array} \begin{array}{c} \begin{array}{ccccc} p=0 & p=1 & \cdots & & p=9 \end{array} \\ \begin{bmatrix} x_{00} & x_{01} & x_{02} & \cdots & x_{09} \\ x_{10} & \vdots & & & \\ x_{20} & \vdots & & & \\ \vdots & \vdots & & & \\ x_{99,0} & x_{99,1} & \cdots & & x_{99,9} \end{bmatrix} \end{array}$$

Simple model

$$y(x_i) = y_i = \tilde{y}_i + \varepsilon_i$$

$$\tilde{y}_i = \beta_0 + \beta \underline{x}_i$$

$$\tilde{y} = X\beta \qquad X \in \mathbb{R}^{n \times p}$$

$$\beta \in \mathbb{R}^{p}$$

$$X_{i*}$$

$$\tilde{y}_i = f(x_i) = \begin{array}{c c} & \begin{array}{cccc} p=0 & p=1 & \cdots & p-1 \end{array} \\ \begin{array}{c} i=0 \\ \vdots \\ i \\ \vdots \\ i=n-1 \end{array} & \begin{bmatrix} x_{00} & x_{01} & \cdots & x_{0p-1} \\ \vdots & & & \\ x_{i0} & & & \\ \vdots & & & \\ x_{n-1} & & & x_{n-1\,p-1} \end{bmatrix} \end{array} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix}$$